The Polar Data Catalogue: Best Practices for Sharing and Archiving Canada's Polar Data

Julie E. Friddell[1*], Gabrielle Alix[1], Yunwei Dong[1], David Friddell[1], Frank Lauritzen[1], and Ellsworth LeDrew[1]

[1*] *Canadian Cryospheric Information Network/Polar Data Catalogue, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2V 3G1, Canada*
Email: julie.friddell@uwaterloo.ca

Summary. Since its online launch in 2007, the Polar Data Catalogue has become Canada's primary online source for Arctic and Antarctic research data and information. Through stable partnerships, the PDC has developed infrastructure and policy for reliable security, discoverability, and access to Canada's polar data and metadata. Recent activities include increased user and community engagement and collaboration, online release of a new bilingual data entry application, registration of DOIs for hundreds of PDC datasets, and conversion of the PDC database to an open format. The PDC is a member of the World Data System and is Canada's National Antarctic Data Centre.

Keywords. Data management, Data access, User engagement, Best practices.

1. Introduction

Scientific research in the polar regions has increased tremendously over the past few decades, with programs such as the ArcticNet Network of Centres of Excellence, Canada's International Polar Year (IPY) programme, the Canadian High Arctic Research Station (CHARS) of Polar Knowledge Canada (POLAR), other government and academic programs, and organizations and communities in northern Canada generating massive amounts of new scientific data and information on Arctic Canada and Antarctica. With these data comes the need to build systems to manage the burgeoning new data resources, to ensure their proper preservation, stewardship, and access. An appropriate data management system must not only respect confidentiality requirements and researchers' rights to publication but also, even more fundamentally, be able to accommodate the vast amounts of data and the huge diversity of topics and fields represented.

2. The Polar Data Catalogue

To help address the data management challenge, ArcticNet, the Canadian IPY programme, Canada's Department of Fisheries and Oceans, Noetix Research Inc., and the Canadian Cryospheric Information Network (CCIN) at the University of Waterloo joined together to develop the Polar Data Catalogue (PDC; https://www.polardata.ca). With additional collaborators Environment and Climate Change Canada; the Beaufort Regional Environmental Assessment, the Northern Contaminants Program, and the Nunavut General Monitoring Plan of Indigenous and Northern Affairs Canada; the international Circumpolar Biodiversity Monitoring Program (CBMP); and numerous others, the PDC has been developed into one of the largest repositories of polar data in Canada. Current holdings include over 2,500 descriptive metadata records of datasets and other polar data resources, almost 2.9 million datafiles, and more than 28,000 satellite images of northern Canada and Antarctic, all of which are available for free download.

3. User Engagement & Collaboration

To enhance management and accessibility of the data in the PDC archive, CCIN staff have actively sought input from partners and users whom we seek to serve, namely researchers and students, northern Canadian community members, and policy and decision makers at all levels of

government and in northern organizations. Recent activities include review of the PDC Lite low-bandwidth data search application (https://www.polardata.ca/pdclite/) by Inuit partners which resulted in a number of interface improvements to better meet user needs. We have also conducted a targeted survey of user requirements for snow and ice data in Canada. In addition to receiving large numbers of specific requests for particular types of snow and ice data and products, the survey demonstrated that people seek raw data as well as interactive visualizations, graphs, and map products which help them to understand the data more quickly and easily.

A significant effort over the last few years has involved hosting three meetings to advance management of polar data within Canada and at the international level. In 2015 and in 2017, CCIN hosted two Canadian Polar Data Workshops (https://secondcanadianpolardataworkshop.wordpress.com) to facilitate collaboration within the growing polar data community in Canada, and, in 2015, co-hosted the Polar Data Forum II (http://www.polar-data-forum.org) to bring the international polar data community together for making progress on priority themes and challenges.

A third engagement activity with three of our partners has resulted in publication of *Data Management Principles and Guidelines for Polar Research and Monitoring in Canada* (https://www.canada.ca/en/polar-knowledge/publications/data-management-principles-and-guidelines-2017-may.html). This document is intended to guide and unify data management requirements and expectations for polar researchers in Canada, streamlining and simplifying work for scientists and for data managers.

4. Website and App Development

To better serve our contributing researchers, some of whom have requested improvements and updates to the functions and usability of the PDC metadata and data Input application, we undertook a major project in 2016 to completely redesign and rebuild our data ingest app. The new application (https://www.polardata.ca/pdcinput/) has a completely modern and improved user interface, uses updated web technologies, is mobile-enabled, and is bilingual in English and French to accommodate the two official languages of Canada. Our next large projects are to redevelop and modernize the CCIN website (https://ccin.ca/) and the PDC Search application, both of which are currently in advanced prototype stages and are expected to be released in 2018.

5. DOIs to Improve Access to Data

Digital Object Identifiers (DOIs) are standard online identifiers that provide long-term links to data, improving the discoverability, accessibility, and citability of the data to which they are assigned. DOIs are viewed by researchers as an incentive to provide their data and thus make depositing their data attractive for researchers. In partnership with the National Research Council, Canada's member of the international DataCite initiative, CCIN has developed a process to assign DOIs and has registered DOIs to over 300 PDC datasets. We can provide DOIs early in the data management process so that the DOI can be recorded in a researcher's journal publications which report on the archived dataset.

6. Conclusions

The collaborative and technical activities over the last couple years have helped CCIN and the PDC grow into a well-respected repository of Arctic and Antarctic data. In recognition of the professionalism of our operation, the PDC has become the 100th member of the World Data System and has been selected as Canada's National Antarctic Data Centre. To continue the progress, we plan to strengthen the dialogue initiated via the Canadian and international Polar Data Workshops and Forum and receive and address additional user needs for improving our operation, so that we can optimally serve our users and contributors in the best possible ways.

Acknowledgments. We would like to thank our partners and supporters who have funded development of the advances reported here as well as numerous additional improvements over the last several years. We also thank the hundreds of researchers who willingly provide their time and data for archiving and sharing.